



Chronic Condition Data Warehouse

Your source for national CMS Medicare and Medicaid research data

CCW User Guide:

Medicare

Risk Score Files

July 2018

Version 1.0

This page intentionally left blank.

TABLE OF CONTENTS

OVERVIEW	1
CHAPTER 1: BACKGROUND	3
CHAPTER 2: CCW PAYMENT RISK SCORE FILES.....	5
A. SOURCE DATA	5
B. CCW BASE RISK ADJUSTMENT FILE	5
C. CCW DETAILED MOF FILES.....	7
1. V12 scores	8
2. V21 scores	9
3. ESRD V21 scores.....	9
4. V22 scores	9
5. RXHCC scores	10
CHAPTER 3: MEDICARE RISK SCORES.....	11
A. CALCULATION OF RISK SCORES	11
B. HIERARCHICAL CONDITION CATEGORIES	13
CHAPTER 4: MEDICARE ENROLLMENT DATA	14
CHAPTER 5: RECEIVING CCW DATA	16
CHAPTER 6: FURTHER ASSISTANCE WITH MEDICARE DATA	17
APPENDIX A: ACRONYM LIST	18
APPENDIX B: COMPARISON OF HCC CONDITIONS – BY MODEL	19

LIST OF TABLES

Table 1. Base File – Resolved Payment Risk Scores	6
Table 2. Model Segment Code Values	6
Table 3. SAS Naming Conventions for HCCs – V12 Example	8
Table 4. SAS Naming Conventions for HCCs – V21 Example	9
Table 5. Risk Adjustment Model Variables and Adjustments.....	11

LIST OF FIGURES

Figure 1. Risk Adjustment Processing to create Research Risk Score files	5
---	---

Overview

Medicare is the primary health insurance program for people age 65 or older, people under age 65 with disabilities, and people of all ages with End-Stage Renal Disease (ESRD). Centers for Medicare & Medicaid Services (CMS) developed risk scores that Medicare uses to adjust payments to Medicare Advantage (MA) plans to account for differences in health status among plan enrollees¹.

CMS uses the Chronic Condition Data Warehouse (CCW) to develop and manage CMS research data resources. The CCW has complete (100 percent) Medicare enrollment and fee-for-service (FFS) claims data, obtained directly from CMS. CCW obtains Medicare Payment Risk Score data files² from CMS. From this source data, CCW has prepared data files to disseminate to researchers who have been approved under a Data Use Agreement (DUA) to obtain Medicare administrative data for research purposes. The CCW Medicare data contain identifiable information, and are subject to the Privacy Act and other Federal government rules and regulations (see the Research Data Assistance Center, or ResDAC, website for information on requesting Medicare data: <http://www.resdac.org/>).

The Medicare Risk Score files are available starting with the 2014 payment year. They are designed to be used with other CCW data products, such as the Master Beneficiary Summary File (MBSF) – which contains a wide range of demographic and Medicare coverage information. When the source data are loaded to the CCW Database, each record is assigned the CCW unique beneficiary identifier (the BENE_ID variable), to facilitate linkage between files – and for longitudinal analyses.

This guide provides users with information that may be helpful in understanding and working with the CCW Risk Score data files.

This document is organized as follows:

[Chapter 1](#) describes the CMS Payment Risk Score Files at a high-level. The two types of files which may be requested are the base (resolved) risk score file and the detailed model output files.

¹ CMS. Medicare, Risk Adjustment webpage. <https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Risk-Adjustors.html>. Accessed 03/19/2018.

² CMS. HCC V22 model, published in “Announcement of Calendar Year (CY) 2014 Medicare Advantage Capitation Rates and Medicare Advantage and Part D Payment Policies and Final Call Letter.” April 2013. <https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/downloads/Announcement2014.pdf>

[Chapter 2](#) describes the layout for the CCW Payment Risk Score files. Note that a full data dictionary is available on ccwdata.org.

[Chapter 3](#) describes the source data for the CMS Payment Risk Score Files, key concepts related to how the CMS Payment Risk Scores are calculated, and how to interpret the scores.

[Chapter 4](#) describes the Medicare enrollment data to use with these data files.

[Chapter 5](#) provides instructions on how to obtain the data files.

[Chapter 6](#) is information regarding where to go with questions.

[Appendix A](#) lists the abbreviations and acronyms used in this document.

[Appendix B](#) identifies the HCC conditions used in the various models.

Chapter 1: Background

Differences in beneficiaries' health status can have a significant influence on health care use and spending. One tool that can help understand those differences are the risk scores that Medicare uses to adjust payments to MA plans to account for differences in health status among plan enrollees. MA plans with enrollees who have higher than average risk scores are expected to incur higher than average costs; therefore, CMS adjusts the capitated payments to MA plans accordingly. Those risk scores are based in part on diagnosis codes that are then grouped into related disease categories known as hierarchical condition categories, or HCCs.

The HCC-based payment risk scores are actuarial estimates regarding how a beneficiaries' fee-for-service (FFS) spending will compare to the overall average for the entire FFS Medicare population.³ The average (normalized) risk score for the overall FFS Medicare population is set at 1.0; beneficiaries with scores greater than that are expected to have above-average spending (and proportionally higher as well – a score of 2.0 means that your costs are expected to be twice the overall average), and vice versa. Risk scores are based on a combination of factors: the beneficiary's age and sex; whether the beneficiary is eligible for Medicaid, first qualified for Medicare on the basis of disability, or lives in an institution (usually a nursing home); and the diagnoses on the beneficiary's claims from the previous year if the beneficiary was enrolled in Part B for all 12 months of the previous year (i.e., the risk score for payment year 2014 is based on diagnoses from claims from 2013). Diagnosis codes are not a model input for beneficiaries with less than 12 months of Part B during the previous year. CMS employs a "new enrollee" model for these beneficiaries.

Each year, CMS evaluates and updates the HCC model to incorporate changes in the diagnosis codes that are mapped in HCCs. Periodically, CMS estimates a new HCC model that is used to calculate a risk score for each beneficiary. In 2014, CMS created what is referred to as "Version 22" of the model.

The basic structure of each risk model is the same – they rely on the same demographic factors and use similar HCCs, but the relative weights they assign to each factor differ. Each beneficiary is assigned a "**resolved**" payment score based on beneficiary's status in a payment month. In general, the status is a function of ESRD status, new enrollee status, and community/institutional status. There are variations on this status based on MA enrollment. Once the risk scores are calculated for each model segment, the scores are then **normalized**. Normalization occurs to create an average risk score of 1.0 for the FFS population for that model. A managed care **coding**

³ Pope GC, Kautter J, Ingber MJ, et al. "Evaluation of the CMS-HCC Risk Adjustment Model." Final Report to CMS. March 2011. Available at http://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Downloads/Evaluation_Risk_Adj_Model_2011.pdf. Accessed 03/19/2018.

intensity factor is also applied to account for differences in the coding patterns (i.e., to account for diagnosis code submission patterns) for Medicare managed care plans (MAs) versus FFS Medicare. Both the normalization and the coding intensity factors can differ by each year and can be found in the HCC annual announcements on the CMS website⁴. Note that the coding intensity adjustment is an adjustment to risk scores for MA payments; however, the resolved risk scores for all beneficiaries include this adjustment, which may not be appropriate for beneficiaries enrolled in FFS.

The CCW makes available two main types of HCC files:

1. **The Base (Resolved) Risk Score File (RAF)** – This is a beneficiary-level data file with 98 variables containing monthly Risk Adjustment Payment Scores. These monthly scores have been normalized and a coding intensity factor has been applied. CMS has determined which of the Part C and Part D segments was most appropriate for each beneficiary/month, and provides the corresponding risk adjustment payment score (i.e., a resolved risk score). Payment scores are included in the file for all model segments (Part C and Part D), even segments not applicable to a beneficiary.

This file is the “base” Risk Score file that all researchers approved to receive CCW Risk Score data will obtain.

2. **Model Output Files (MOFs)** - There are files for the five (5) HCC models for payment year 2014. Each is a beneficiary-level file that contains the model-specific “raw” risk scores. The raw risk scores are not normalized and have not had a coding intensity factor applied, nor are they “resolved” to identify the appropriate HCC model applicable to the beneficiary for the month. Each file also includes the HCC conditions required for the model. The detailed MOFs are:
 - a. V12 scores – 169 variables
 - b. V21 scores – 237 variables
 - c. V22 scores – 200 variables
 - d. ESRD V21 scores – 265 variables
 - e. RXHCC scores – 204 variables

Investigators who wish to obtain the detailed MOFs are required to provide additional justification (beyond what is required for the Base file).

These files are described in greater detail in the next Chapter.

⁴CMS. Medicare Advantage Rates & Statistics. Announcements and Documents (updated each year) <https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Announcements-and-Documents.html>

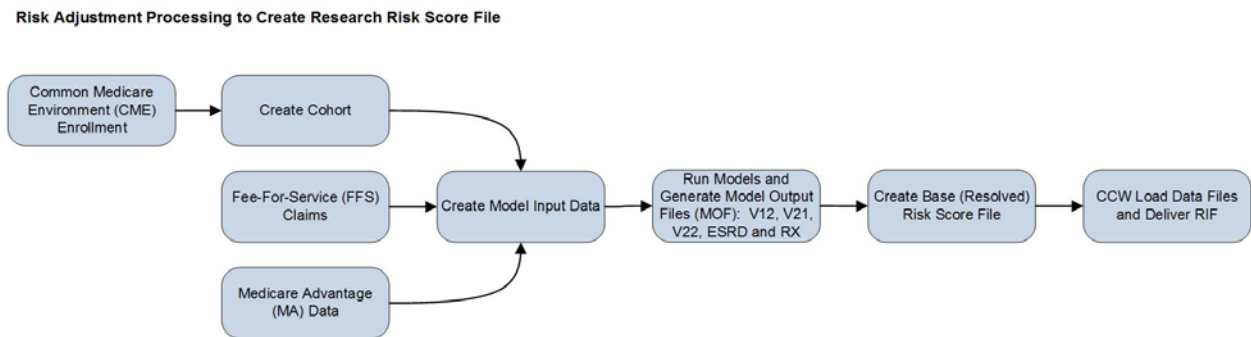
Chapter 2: CCW Payment Risk Score Files

The two types of CCW HCC files are described in greater detail within this Chapter.

A. Source Data

Multiple sources are used to create the input data to produce the model risk scores. These sources are the CMS Common Medicare Environment (CME) tables for the enrollment and demographic data, CMS fee-for-service (FFS) claims, and MA plan-submitted diagnosis data. The final model runs for a given payment year use current beneficiary demographic information and diagnosis data for the data collection period submitted by January after the payment year (i.e. January 2015 for payment year 2014, covering services provided in 2013). A visual depiction of the CMS data sources and processing to create the Risk Score Files for researchers is in Figure 1.

Figure 1. Risk Adjustment Processing to create Research Risk Score files



The risk adjustment models use a prospective, rather than a concurrent model. That means that information from the year prior to the payment year (or prediction year) is used to identify the HCCs from claims. For the 2014 CCW HCC data files, the payment/prediction year is 2014 – and information from Medicare services in 2013 is used in the model.

CMS runs the risk adjustment model process at least three times for each payment year. The CCW Base and MOF files represent the final model run for each payment year. The risk adjustment model uses standard algorithms to generate the raw and payment score files.

Each beneficiary is run through each model segment for each model. Using these various model output files, a single record per beneficiary is created, with the risk scores created for the beneficiary normalized, adjusted and arrayed on a single record.

B. CCW Base Risk Adjustment File

The base (resolved) risk score file provided by the CCW to researchers uses the payment algorithms to identify the beneficiary’s correct monthly payment score, i.e. the monthly

“resolved” risk scores for each beneficiary⁵. There are four types of variables in the file: 1) overall monthly scores, 2) monthly model indicators, 3) long-term monthly institutional indicators, and 4) non-resolved payment scores. Payment risk scores are calculated for all model segments for a beneficiary, regardless of whether or not the model segment actually applies to the beneficiary. Throughout this document, when we identify a particular data variable by name we will often identify the specific SAS name which will appear in all capitals.

Table 1. Base File – Resolved Payment Risk Scores

Variable	Label
BENE_ID	CCW Encrypted Beneficiary ID
PTC_RISK_SCORE_01 - 12	Monthly Part C risk score 01 (January) – 12 (December)
PTC_MODEL_SGMT_CD_01 - 12	Monthly Part C Model Segment Code 01 (January) – 12 (December)
PTD_RISK_SCORE_01 - 12	Monthly Part D risk score 01 (January) – 12 (December)
PTD_MODEL_SGMT_CD_01 - 12	Monthly Part C Model Segment Code 01 (January) – 12 (December)
PTD_RISK_SCORE_01 - 12	Monthly Part D risk score 01 (January) – 12 (December)

The Part C Model Segment Code is used to indicate which of the HCC model segments is applied to the beneficiary for each month. Values are listed in the table below.

Table 2. Model Segment Code Values

PTC_MODEL_SGMT_CD_MM	Description
C	Community beneficiary
C1	Community beneficiary with a kidney transplant in past 4-9 months (ESRD)
C2	Community beneficiary with a kidney transplant 10+ months ago (ESRD)
D	Beneficiary receiving dialysis (ESRD)
E	New enrollee
ED	New enrollee receiving dialysis (ESRD)
E1	New enrollee with a kidney transplant in past 4-9 months (ESRD)
E2	New enrollee with a kidney transplant 10+ months ago (ESRD)
G1	Beneficiary with a kidney transplant in current month (ESRD)
G2	Beneficiary with a kidney transplant in past 2-3 months (ESRD)
I	Institutional beneficiary
I1	Institutional with a kidney transplant in past 4-9 months
I2	Institutional beneficiary with a kidney transplant 10+ months ago
SE	New Enrollee Chronic Care Special Needs Plan (SNP)
Null/missing	Not applicable

Any code other than C, E, SE, or I indicates some type of ESRD status.

⁵ Note that it may not be possible to replicate the actual payment risk score for a plan using these files due to changes in the status of the beneficiary.

The “resolution” process determines which segment actually applies to the beneficiary for the month. The Base (resolved) risk score file include monthly indicators that a beneficiary was identified as long-term institutional for that month (LTI; i.e., institutional indicator). The demographic variables that are in the Base file are demographics that were used in the risk score development, and are obtained from CMS within the Risk Score file; it is possible that some values may not exactly match values in the CCW Medicare Beneficiary Summary File (MBSF) due to timing differences. Additional details regarding the MBSF appear later in this document (Chapter 4).

The largest proportion of individuals live in the community setting. Individuals in long term care facilities have different medical needs which can be associated with greater payment risk.

As with the Medicare Part C HCCs, several Part D model segments (also referred to as Rx model segments) have been developed. These include: (1) aged, (2) disabled, (3) institutional and (4) new enrollee. Some of these segments are further divided into low income and not low income. As with Part C risk scores, Rx risk scores are calculated using FFS claims and MA reported diagnosis information to predict expenditures associated with prescription drugs. Additional details are available on the CMS website⁶⁷.

A listing of all variables included in the data file appears in the CCW Data Dictionary (see <https://www.ccwdata.org/web/guest/data-dictionaries>).

It is possible that some investigators may wish to calculate a single annual payment risk score. If this is desirable, then CCW recommends averaging the monthly scores to calculate the annual score.

C. CCW Detailed MOF Files

We envision that most investigators will be able to conduct their research using just the Base file since all of the applicable resolved beneficiary scores appear in this file. However, a brief description of the Detailed MOFs follows. CMS and CCW do not provide technical assistance or user support to investigators who wish to use these files.

For 2014 payment year, there are five MOFs. The MOF files are model specific, and contain the risk scores that potentially can be assigned to a beneficiary for payment. For example, the ESRD

⁶CMS. Annual “Rate Announcement and Call Letters” are available for Prescription Drug Coverage Contractors. See, for example: <https://www.cms.gov/Medicare/Prescription-Drug-Coverage/PrescriptionDrugCovContra/index.html>

⁷ CMS. Annual “Rate Announcement and Call Letters” are available for Prescription Drug Coverage Contractors. See, for example: <https://www.cms.gov/Medicare/Prescription-Drug-Coverage/PrescriptionDrugCovContra/index.html>

model contains the various segments that are used to calculate the risk score for the set of ESRD statuses assigned to beneficiary: dialysis, transplant, and post-graft. The V21 model is used for PACE enrollees, and the V12 and V22 models are used to create a blended risk score for MA enrollees.

The MOFs for each of the prospective risk adjustment models include only the HCCs that best predict future (next year's, i.e., the payment year [2014 for the CCW files]) Medicare Part A and Part B expenditures. Model versions are updated for a variety of reasons, including changes in valid diagnoses mapping to the HCCs, updates to accommodate more recent years of data, as recalibrations to incorporate clinical and other updates. Not all model versions are used for payment.⁸

A listing of the HCCs used for different versions of the MOF are included in Appendix B: Comparison of HCC Conditions – by Model. CCW provides unique SAS names for each variable in the MOFs, so that it is easy to trace which version of the HCCs is represented in the data. If a beneficiary has more than one condition within a particular HCC hierarchy, only the most severe condition is used when determining his or her risk score. Each condition variable has a value of 1 if the beneficiary had the condition in the prior year; otherwise the value is 0 (zero).

1. V12 scores

Version 12 (V12) of the HCC model is available through 2014. There are a total of 189 distinct condition categories represented in the V12 of the HCC risk-adjustment model, which are then grouped into 70 disease hierarchies (see listing in Appendix B). CCW provides unique SAS names for each variable in the MOFs; all of the HCCs within the V12 file have a “V12_” prefix. For example, HCC1 (HIV/AIDS) has the SAS name V12_HCC1 in this data file (Table 3).

Table 3. SAS Naming Conventions for HCCs – V12 Example

SAS Variable Name	Condition
V12_HCC1	HIV / AIDS
V12_HCC2	Septicemia / shock
V12_HCC5	Opportunistic infections
V12_HCC7	Metastatic cancer and acute leukemia

Each variable has a value of 1 if the beneficiary had the condition in the prior year; otherwise the value is 0 (zero).

⁸ Pope GC, Kautter J, Ingber MJ, et al. “Evaluation of the CMS-HCC Risk Adjustment Model.” Final Report to CMS. March 2011. Available at http://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Downloads/Evaluation_Risk_Adj_Model_2011.pdf

There are also variables to indicate some disease/disease combinations, and some disability/disease combinations. Demographic variables that are related to the segment scores (e.g., gender, age groupings and original reason for enrollment in Medicare) are also available.

2. V21 scores

In 2012, there was an update in the HCC model called version 21 (V21). In 2014 this model is used for PACE and ESRD enrollees. This updated model incorporated additional HCC conditions, different disease interactions, and different disabled HCC interactions terms to calculate a HCC score compared to the previous model⁹. There are 237 variables in the file, of which 87 are variables for conditions that are included in the V21 MOF (see table in Appendix B). All the HCCs within the V21 file have a SAS variable name with “V21_” prefix (see Table 4. For example, HCC1 (HIV/AIDS) has the SAS name V21_HCC1 in this data file (whereas the same HCC in the V12 file was called V12_HCC1). Note that **the HCC numbering applies to different condition categories in different models**. For example in V12 HCC5 was Opportunistic Infections (V12_HCC5); for V21 HCC6 is for Opportunistic Infections (V21_HCC6).

Table 4. SAS Naming Conventions for HCCs – V21 Example

SAS Variable Name	Condition
V21_HCC1	HIV/AIDS (V21 model)
V21_HCC2	Septicemia/Shock (V21 HCC2)
V21_HCC6	Opportunistic Infections (V21 model)
V21_HCC8	Metastatic Cancer/Acute Leukemia (V21 model)

3. ESRD V21 scores

Individuals who experience End Stage Renal Disease (ESRD) have high medical costs and more extensive medical needs than the general Medicare population; therefore, this group of individuals have their own HCC model. The ESRD HCC model divides the population into 4segments: (1) those undergoing dialysis, (2) those who have received kidney transplants (from the time of transplant to 3 months post graft), (3) those 4 to 9 months after a transplant (a functioning graft), and (4) those who are 10 months or more after a transplant. There are 265 variables in the ESRD V21 file. This file contains the same 87 HCC conditions as the V21 file.

4. V22 scores

Beginning in 2014, a new HCC model was incorporated into the risk adjustment model. There are a total of 200 variables in the V22 file. Of these, 79 variables represent distinct condition

⁹ Pope GC, Kautter J, Ingber MJ, et al. “Evaluation of the CMS-HCC Risk Adjustment Model.” Final Report to CMS. March 2011. Available at http://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Downloads/Evaluation_Risk_Adj_Model_2011.pdf

categories in version 22 of the HCC risk-adjustment model¹⁰, which are then grouped into 31 disease hierarchies (see Appendix B). The SAS variable names for each of these HCCs begins with “V22_”.

5. RXHCC scores

The Rx model is used to estimate prescription drug use in the Medicare population. As a result there are similar characteristics as the other HCC models. Diagnosis codes are used to calculate Rx HCC categories; however they use a different selection of HCCs to calculate a risk score than is used for the other Part C HCCs. In the Rx HCC model, a total of 78 HCC condition categories are included in the calculation of a risk score. The conditions included in the model are clinically meaningful with regard to expected prescription drug expenditures. For example, hypertension is included in the Rx HCC score calculation but not included in the V21/V12 models. There are 204 variables in the RX HCC file.

¹⁰ CMS. “HHS-Operated Risk Adjustment Methodology Meeting. Discussion Paper.” March 24, 2016. <https://www.cms.gov/CCIIO/Resources/Forms-Reports-and-Other-Resources/Downloads/RA-March-31-White-Paper-032416.pdf> Accessed 03/19/2018.

Chapter 3: Medicare Risk Scores

A. Calculation of Risk Scores

The purpose of the Medicare risk scores is to estimate a relative cost factor. (i.e., it is a payment risk score). CMS calculates individual beneficiary-level risk scores by adding the relative factors associated with each beneficiary’s demographic and disease factors. The CMS Payment Risk Score is built up each year.¹¹ The initial score for each month is set to 0, then the weight/coefficient for each applicable factor for the beneficiary is added (i.e., the monthly score is the sum of all risk score factors in each model for the beneficiary).

The heart of the risk adjustment model is the clinical classification system that organizes the thousands of International Classification of Diseases (ICD) diagnosis codes into a coherent system of diagnostic categories.¹²The diagnostic classification is key in determining the ability of a risk adjustment model to distinguish high from low cost individuals. CMS incorporates information regarding whether the beneficiary appeared to be treated for particular conditions; these condition categories appear in the MOFs (refer to CCW Data dictionary and also the listing of HCCs for each model that appears in Appendix B). The conditions are hierarchical, which means that the most severe within a category is selected; scores from all unrelated conditions categories are accumulated. Not all claims trigger a condition that is meaningful for the HCCs (some combinations of conditions are important – and the effect of the combination is not just additive).

A visual depiction of the categories of data elements that are used in determining the resolved risk scores are found in Figure 2.

Table 5. Risk Adjustment Model Variables and Adjustments

Demographic Variables: Age/Sex Originally entitled to Medicare due to disability Medicaid enrollment status	There are relative factors associated with each demographic variable.
Disease Variables: Disease Hierarchical Condition Categories (HCCs) Disease/Disabled interactions	CMS uses diagnoses submitted by plans and fee-for-service providers to assign HCCs and interactions for each beneficiary. There are relative factors associated with each HCC and interaction.
Sum of Factors. Demographic + Disease = Raw Risk Scores	

¹¹ Information regarding the Risk Model announcements appear on the CMS webpage. See, for example: <https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Announcements-and-Documents.html> . Accessed 06/20/2018.

¹² CMS. “HHS-Operated Risk Adjustment Methodology Meeting. Discussion Paper.” March 24, 2016. <https://www.cms.gov/CCIIO/Resources/Forms-Reports-and-Other-Resources/Downloads/RA-March-31-White-Paper-032416.pdf> Accessed 03/19/2018.

The relative factors for all of the demographic variables, HCCs and interactions are added together. The result is the raw risk score.	
Payment Adjustments:	A normalization factor is applied to keep the average FFS risk score at 1.0
	A coding pattern adjustment is applied to account for differential coding patterns between MA plans and FFS. Note that this adjustment is applied to all scores (not limited to scores for MA enrollees).
The final step is to resolve the risk score for each beneficiary.	

CMS applies adjustments to the raw risk scores to convert them to payment scores:

- **Normalization.** The normalization factor is used to keep the FFS average risk score to be equal to 1.0 for the payment year. Due to the lag between the time the model is estimated and converted to a relative risk score, and the payment year, changes in the demographic composition and health conditions of the FFS cohort may result in an average risk score not equal to 1.0. The normalization factor is designed to correct for this situation.
- **Coding intensity.** The adjustment for MA coding pattern differences (i.e., the MA coding intensity adjustment) is designed to account for the relative difference in reporting of diagnosis codes between the FFS population used to estimate the models and the MA cohort that is being reimbursed based on the model risk score.

Note: The adjustment for MA coding pattern differences is already factored in to the resolved risk scores for all beneficiaries – even those who were not enrolled in MA during the year. This is because the objective of the CMS Risk Scores is to determine MA payment scores. As a result, the scores for FFS beneficiaries are assigned an artificially low risk score (i.e., since the coding intensity adjustment is $0.95 * \text{the resolved risk score}$). Investigators may “back-out” this coding intensity adjustment by identifying FFS beneficiaries (i.e., using the CCW MBSF) and dividing the resolved risk scores by 0.95. The actual adjustment factor may vary by year – and is published in the CMS Annual Rate Announcement and Call Letter for MA plans.
- **Resolved Scores.** The last step in the risk adjustment process is to determine which of the beneficiary’s converted payment scores applies to the beneficiary for a given month. This resolved risk score is what determines the dollar impact on the payment to the MA plan.

The Base(resolved) risk score file has each beneficiary’s monthly “resolved” risk scores (i.e., the PTC_RISK_SCORE_MM variable, where MM is the month 01 through 12), and monthly HCC Model Segment Codes (the risk score factor code, or PTC_MODEL_SGMT_CD_MM) that indicates which segment was used to calculate the risk score for that particular month. This Base file also contains

the converted payment scores for the beneficiary that were not determined to apply for a given month.

B. Hierarchical Condition Categories

The CMS-HCC diagnostic classification system begins by classifying all ICD-9-CM diagnosis codes (and beginning October 1, 2015, ICD-10-CM codes) into Diagnostic Groups. Each ICD-9-CM code maps to exactly one Diagnostic Group, which represents a well-specified medical condition or set of conditions (note that a small number of diagnosis codes mapped into more than one HCC). The Diagnostic Groupings are further aggregated into Condition Categories (CCs). CCs describe a broader set of similar diseases. Although they are not as homogeneous as the Diagnostic Groupings, diseases within a CC are related clinically and with respect to cost. An example is the CC for *Diabetes with Acute Complications*, which includes in addition to the Diagnostic Group for *Type II Diabetes with Ketoacidosis or Coma*, also the Diagnostic Group for *Type I Diabetes* and *Secondary Diabetes* (each with ketoacidosis or coma).

Hierarchies are imposed among related CCs, so that a person is coded for only the most severe manifestation among related diseases. After imposing hierarchies, CCs become Hierarchical Condition Categories (HCCs). For example, diabetes diagnosis codes are organized in the Diabetes hierarchy, consisting of three CCs arranged in descending order of clinical severity and cost, from 1) *Diabetes with Acute Complications* to 2) *Diabetes with Chronic Complications* to 3) *Diabetes without Complication*. Thus, a person with diagnosis code of *Diabetes with Acute Complications* is excluded from being coded with *Diabetes with Chronic Complications* and is also excluded from being coded with *Diabetes without Complication*. Similarly, a person with a diagnosis code of *Diabetes with Chronic Complications* is excluded from being coded with *Diabetes without Complication*. Although HCCs reflect hierarchies among related disease categories, for unrelated diseases, HCCs accumulate, i.e., the model is “additive.” For example, a female with both *Rheumatoid Arthritis* and *Breast Cancer* has (at least) two separate HCCs coded, and her predicted cost will reflect increments for both conditions.

Because a single individual may be coded for none, one, or more than one HCC, the CMS-HCC model can individually price tens of thousands of distinct clinical profiles. The model’s structure thus provides, and predicts from, a detailed comprehensive clinical profile for each individual. An evaluation of the CMS Risk Models is publicly available on the CMS website¹³.

¹³ Pope GC, Kautter J, Ingber MJ, et al. “Evaluation of the CMS-HCC Risk Adjustment Model.” Final Report to CMS. March 2011. Available at “ https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/downloads/Evaluation_Risk_Adj_Model_2011.pdf ”

Chapter 4: Medicare Enrollment Data

Investigators using the CCW Risk Score Files will want to use information from Medicare Enrollment files to identify and describe the study population. Data for 100% of Medicare-enrolled beneficiaries is available from the CCW. The CCW updates the annual enrollment data files known as the Master Beneficiary Summary File (MBSF) using downloads from the CMS Common Medicare Environment (CME) database. Each annual file includes all beneficiaries documented as being alive for some part of the calendar year and enrolled in the Medicare program (Part A or Part B) for at least one month of the year.

Upon enrollment in Medicare, each beneficiary receives a health insurance claim (HIC) number. HIC numbers appear on Medicare claims and encounter data and were traditionally used to identify beneficiaries, but they are not ideal for research purposes because beneficiaries may have more than one HIC number during their lives due to events such as the death of a spouse or remarriage. The CCW creates a unique beneficiary identifier (known as the BENE_ID) using information from the CMS enrollment database, which contains a variety of identifying information for each beneficiary, such as his or her HIC, sex, Social Security number (SSN), and date of birth. The BENE_ID uniquely identifies each beneficiary, and is used in CCW data files in lieu of more sensitive identifiers such as the HIC or SSN (unless otherwise specified/approved in the CMS data use agreement [DUA]). The BENE_ID allows linkage of an individual's data across data sources/types.

The unique CCW beneficiary identifier field is specific to the CCW and is not applicable to any other identification system or data source. This identifier is encrypted prior to delivering the data files to researchers. In addition, all data files delivered to researchers are encrypted (see Encryption Information in [Chapter 6](#) for details). Each research request employs a different encryption key for the beneficiary identifier field and the data files.

The record layouts for all of the CCW data files may be obtained from the [ccwdata.org](https://www.ccwdata.org/web/guest/data-dictionaries) website, under the Data Dictionaries tab (<https://www.ccwdata.org/web/guest/data-dictionaries>).

The MBSF contains information on beneficiaries' demographic characteristics and details of their enrollment in Medicare.

Examples of the types of information in the MBSF include:

- **Demographic** – age, race, sex, date of death;
- **Geographic** – state, county, zip code;

- **Enrollment** – the start date for Medicare coverage; how the beneficiary is entitled to Medicare (both the original reason and the current reason, which can differ); and monthly information on eligibility (Part A, Part B, or both), enrollment in MA (aka Medicare Part C), and enrollment in Part D. Please note nearly 100% of beneficiaries enrolled in MA have coverage under both Part A and Part B. In contrast, about 88% of beneficiaries enrolled in FFS have coverage under both Part A and Part B.

The Base Beneficiary Summary File is also known as the MBSF Part A, B, C and D segment.

Although there is a small amount of beneficiary demographic information on the MOFs, in general we recommend that you use the MBSF to obtain these data. The MOFs were created at a different point in time, during which actual values (e.g., age or disability status) may have changed, and in rare instances used different decision rules for populating fields than the MBSF. The MBSF should be considered the definitive source of Medicare beneficiary enrollment and demographic information.

Chapter 5: Receiving CCW Data

The CCW Medicare RIFs are provided to academic researchers and certain government agencies, to conduct approved research studies under a Data Use Agreement (DUA). The CCW Medicare data contain identifiable information, and are subject to the Privacy Act and other Federal government rules and regulations (see ResDAC website for information on requesting Medicare data <http://www.resdac.org/>).

Once the DUA is in place and payment for the files has been received, data files are shipped to the requestor on either a USB external hard drive or a DVD/CD. These data files are packaged as encrypted self-decrypting archive (SDA) files (see the CCW User Guide at ccwdata.org for additional information on encryption). The decryption password is sent electronically via email. When you receive the data package (via hard drive or DVD/CD), copy them from the shipping media to your local workspace. Note some data shipped on a hard drive can be decrypted on that hard drive, depending on the size of the data files. Using the password provided to you via email, follow the Decryption Instructions enclosed in the data package. Each SDA contains the data file(s), SAS[®] code and a file transfer summary (.fts) file which can be used to verify the data was read in correctly.

Chapter 6: Further Assistance with Medicare Data

Additional information regarding the Medicare Advantage rates and statistics including annual HCC announcements can be found on the CMS webpage.¹⁴

ResDAC offers free assistance to those using Medicare data for research. The ResDAC website provides links to descriptions of the CMS data available, request procedures, supporting documentation, workshops on how to use Medicare data and other helpful resources. Visit the ResDAC website at (<http://www.resdac.org>) for additional information.

ResDAC is a CMS contractor and requests for assistance in the application, obtaining, or using the CCW data should first be submitted to ResDAC. Investigators can reach ResDAC by phone at 1-888-973-7322, e-mail at resdac@umn.edu, or online at (<http://www.resdac.org>).

In the event that a ResDAC technical advisor is not able to answer the question, the technical advisor will direct the researcher to the appropriate person. If additional CMS data (data not available from the CCW) is required to meet research objectives, or the researcher has any questions about other data sources, the researcher can review all available CMS data by visiting the ResDAC website and contact ResDAC for further assistance.

www.ccwdata.org

Email: CMSdata@gdit.com

Phone: 1-866-766-1915

¹⁴ For example: <https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Announcements-and-Documents.html>

Appendix A: Acronym List

Acronym	Definition
CC	Condition Category
CCW	Chronic Conditions Data Warehouse
CME	Common Medicare Environment
CMS	Centers for Medicare & Medicaid Services
DUA	Data Use Agreement
ESRD	End stage renal disease
FFS	Fee-For-Service
HCC	Hierarchical Condition Category
HIC	Health Insurance Claim Number
ICD	International Classification of Diseases
LTI	Long term institutional
MA	Medicare Advantage
MBSF	Master Beneficiary Summary File
MOF	Model Output File
PACE	Program of All-Inclusive Care for the Elderly
RAF	Risk Adjustment Factor
RIF	Research Identifiable File

Appendix B: Comparison of HCC Conditions – by Model

The HCC numbering applies to different condition categories in different models (see Table below¹⁵). For example in V12 HCC 161 was Traumatic Amputation; for V22 HCC 161 is for Chronic Ulcer of Skin, Except Pressure Ulcers. If a beneficiary has more than one condition within a particular hierarchy, only the most severe condition is used when determining his or her risk score. For example, the end-stage liver disease hierarchy has three different diseases that are subsumed by the HCC when one of the following diseases is also present: cirrhosis, chronic hepatitis, and Coma, Brain Compression/Anoxic Damage. Each condition has a weight that reflects its marginal contribution to a beneficiary’s total expected Medicare costs (weights not shown in Table; refer to CMS “Announcement” letters¹⁶). Each variable has a value of 1 if the beneficiary had the condition in the prior year; otherwise the value is 0 (zero).

V12 Model (with 70 HCCs)		V21 Model (with 83 HCCs)		V22 Model (with 79 HCCs)		Category Short Name
HCC	Description	HCC	Description	HCC	Description	
HCC1	HIV/AIDS	HCC1	HIV/AIDS	HCC1	HIV/AIDS	Infection
HCC2	Septicemia/Shock	HCC2	Septicemia, Sepsis, Systemic Inflammatory Response Syndrome/Shock	HCC2	Septicemia, Sepsis, Systemic Inflammatory Response Syndrome/Shock	
HCC5	Opportunistic Infections	HCC6	Opportunistic Infections	HCC6	Opportunistic Infections	
HCC7	Metastatic Cancer and Acute Leukemia	HCC8	Metastatic Cancer and Acute Leukemia	HCC8	Metastatic Cancer and Acute Leukemia	Neoplasm
HCC8	Lung, Upper Digestive Tract, and Other Severe Cancers	HCC9	Lung and Other Severe Cancers	HCC9	Lung and Other Severe Cancers	
HCC9	Lymphatic, Head and Neck, Brain, and Other Major Cancers	HCC10	Lymphoma and Other Cancers	HCC10	Lymphoma and Other Cancers	
HCC10	Breast, Prostate, Colorectal and Other Cancers and Tumors	HCC11	Colorectal, Bladder, and Other Cancers	HCC11	Colorectal, Bladder, and Other Cancers	

¹⁵ Table is adapted from Table 5, CMS “Advance Notice of Methodological Changes for Calendar Year (CY) 2014 for Medicare Advantage (MA) Capitation Rates, Part C and Part D Payment Policies and 2014 Call Letter.” Downloaded from: <https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Downloads/Advance2014.pdf>

¹⁶ CMS. Medicare Health Plans. Medicare Advantage Advance Notice or Call Letter. For example: <https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Downloads/Advance2019Part1.pdf>

V12 Model (with 70 HCCs)		V21 Model (with 83 HCCs)		V22 Model (with 79 HCCs)		Category Short Name
HCC	Description	HCC	Description	HCC	Description	
		HCC12	Breast, Prostate, and Other Cancers and Tumors	HCC12	Breast, Prostate, and Other Cancers and Tumors	
HCC15	Diabetes with Renal or Peripheral Circulatory Manifestation	HCC17	Diabetes with Acute Complications	HCC17	Diabetes with Acute Complications	Diabetes
HCC16	Diabetes with Neurologic or Other Specified Manifestation	HCC18	Diabetes with Chronic Complications	HCC18	Diabetes with Chronic Complications	
HCC17	Diabetes with Acute Complications	HCC19	Diabetes without Complication	HCC19	Diabetes without Complication	
HCC18	Diabetes with Ophthalmologic or Unspecified Manifestation					
HCC19	Diabetes without Complication					
HCC21	Protein-Calorie Malnutrition	HCC21	Protein-Calorie Malnutrition	HCC21	Protein-Calorie Malnutrition	Metabolic
		HCC22	Morbid Obesity	HCC22	Morbid Obesity	
		HCC23	Other Significant Endocrine and Metabolic Disorders	HCC23	Other Significant Endocrine and Metabolic Disorders	
HCC25	End-Stage Liver Disease	HCC27	End-Stage Liver Disease	HCC27	End-Stage Liver Disease	Liver
HCC26	Cirrhosis of Liver	HCC28	Cirrhosis of Liver	HCC28	Cirrhosis of Liver	
HCC27	Chronic Hepatitis	HCC29	Chronic Hepatitis	HCC29	Chronic Hepatitis	
HCC31	Intestinal Obstruction/Perforation Pancreatic Disease	HCC33	Intestinal Obstruction/Perforation	HCC33	Intestinal Obstruction/Perforation	Gastrointestinal
HCC32		HCC34	Chronic Pancreatitis	HCC34	Chronic Pancreatitis	
HCC33	Inflammatory Bowel Disease	HCC35	Inflammatory Bowel Disease	HCC35	Inflammatory Bowel Disease	
HCC37	Bone/Joint/Muscle Infections/Necrosis	HCC39	Bone/Joint/Muscle Infections/Necrosis	HCC39	Bone/Joint/Muscle Infections/Necrosis	Musculoskeletal
HCC38	Rheumatoid Arthritis and Inflammatory Connective Tissue Disease	HCC40	Rheumatoid Arthritis and Inflammatory Connective Tissue Disease	HCC40	Rheumatoid Arthritis and Inflammatory Connective Tissue Disease	
HCC44	Severe Hematological Disorders	HCC46	Severe Hematological Disorders	HCC46	Severe Hematological Disorders	Blood
HCC45	Disorders of Immunity	HCC47	Disorders of Immunity	HCC47	Disorders of Immunity	

V12 Model (with 70 HCCs)		V21 Model (with 83 HCCs)		V22 Model (with 79 HCCs)		Category Short Name	
HCC	Description	HCC	Description	HCC	Description		
		HCC48	Coagulation Defects and Other Specified Hematological Disorders	HCC48	Coagulation Defects and Other Specified Hematological Disorders		
HCC51	Drug/Alcohol Psychosis	HCC54	Drug/Alcohol Psychosis	HCC54	Drug/Alcohol Psychosis	Substance Abuse	
HCC52	Drug/Alcohol Dependence	HCC55	Drug/Alcohol Dependence	HCC55	Drug/Alcohol Dependence		
HCC54	Schizophrenia	HCC57	Schizophrenia	HCC57	Schizophrenia	Psychiatric	
HCC55	Major Depressive, Bipolar, and Paranoid Disorders	HCC58	Major Depressive, Bipolar, and Paranoid Disorders	HCC58	Major Depressive, Bipolar, and Paranoid Disorders		
HCC67	Quadriplegia, Other Extensive Paralysis	HCC70	Quadriplegia	HCC70	Quadriplegia	Spinal	
HCC68	Paraplegia	HCC71	Paraplegia	HCC71	Paraplegia		
HCC69	Spinal Cord Disorders/Injuries	HCC72	Spinal Cord Disorders/Injuries	HCC72	Spinal Cord Disorders/Injuries		
HCC70	Muscular Dystrophy	HCC73	Amyotrophic Lateral Sclerosis and Other Motor Neuron Disease	HCC73	Amyotrophic Lateral Sclerosis and Other Motor Neuron Disease	Neurological	
HCC71	Polyneuropathy	HCC74	Cerebral Palsy	HCC74	Cerebral Palsy		
HCC72	Multiple Sclerosis	HCC75	Polyneuropathy	HCC75	Myasthenia Gravis/Myoneural Disorders and Guillain-Barre Syndrome/Inflammatory and Toxic Neuropathy		
HCC73	Parkinson's and Huntington's Diseases	HCC76	Muscular Dystrophy	HCC76	Muscular Dystrophy		
HCC74	Seizure Disorders and Convulsions	HCC77	Multiple Sclerosis	HCC77	Multiple Sclerosis		
HCC75	Coma, Brain Compression/Anoxic Damage	HCC78	Parkinson's and Huntington's Diseases	HCC78	Parkinson's and Huntington's Diseases		
		HCC79	Seizure Disorders and Convulsions	HCC79	Seizure Disorders and Convulsions		
		HCC80	Coma, Brain Compression/Anoxic Damage	HCC80	Coma, Brain Compression/Anoxic Damage		
HCC77	Respirator Dependence/Tracheostomy Status	HCC82	Respirator Dependence/Tracheostomy Status	HCC82	Respirator Dependence/Tracheostomy Status		Arrest
HCC78	Respiratory Arrest	HCC83	Respiratory Arrest	HCC83	Respiratory Arrest		

V12 Model (with 70 HCCs)		V21 Model (with 83 HCCs)		V22 Model (with 79 HCCs)		Category Short Name
HCC	Description	HCC	Description	HCC	Description	
HCC79	Cardio-Respiratory Failure and Shock	HCC84	Cardio-Respiratory Failure and Shock	HCC84	Cardio-Respiratory Failure and Shock	
HCC80	Congestive Heart Failure	HCC85	Congestive Heart Failure	HCC85	Congestive Heart Failure	Heart
HCC81	Acute Myocardial Infarction	HCC86	Acute Myocardial Infarction	HCC86	Acute Myocardial Infarction	
HCC82	Unstable Angina and Other Acute Ischemic Heart Disease	HCC87	Unstable Angina and Other Acute Ischemic Heart Disease	HCC87	Unstable Angina and Other Acute Ischemic Heart Disease	
HCC83	Angina Pectoris/Old Myocardial Infraction	HCC88	Angina Pectoris	HCC88	Angina Pectoris	
HCC92	Specified Heart Arrhythmias	HCC96	Specified Heart Arrhythmias	HCC96	Specified Heart Arrhythmias	
HCC95	Cerebral Hemorrhage	HCC99	Cerebral Hemorrhage	HCC99	Cerebral Hemorrhage	
HCC96	Ischemic or Unspecified Stroke	HCC100	Ischemic or Unspecified Stroke	HCC100	Ischemic or Unspecified Stroke	
HCC100	Hemiplegia/Hemiparesis	HCC103	Hemiplegia/Hemiparesis	HCC103	Hemiplegia/Hemiparesis	
HCC101	Cerebral Palsy and Other Paralytic Syndromes	HCC104	Monoplegia, Other Paralytic Syndromes	HCC104	Monoplegia, Other Paralytic Syndromes	
HCC104	Vascular Disease with Complications	HCC106	Atherosclerosis of the Extremities with Ulceration or Gangrene	HCC106	Atherosclerosis of the Extremities with Ulceration or Gangrene	Vascular
HCC105	Vascular Disease	HCC107	Vascular Disease with Complications	HCC107	Vascular Disease with Complications	
		HCC108	Vascular Disease	HCC108	Vascular Disease	
HCC107	Cystic Fibrosis	HCC110	Cystic Fibrosis	HCC110	Cystic Fibrosis	Lung
HCC108	Chronic Obstructive Pulmonary Disease	HCC111	Chronic Obstructive Pulmonary Disease	HCC111	Chronic Obstructive Pulmonary Disease	
HCC111	Aspiration and Specified Bacterial Pneumonias	HCC112	Fibrosis of Lung and Other Chronic Lung Disorders	HCC112	Fibrosis of Lung and Other Chronic Lung Disorders	
HCC112	Pneumococcal Pneumonia, Empyema, Lung Abscess	HCC114	Aspiration and Specified Bacterial Pneumonias	HCC114	Aspiration and Specified Bacterial Pneumonias	
		HCC115	Pneumococcal Pneumonia, Empyema, Lung Abscess	HCC115	Pneumococcal Pneumonia, Empyema, Lung Abscess	
HCC119	Proliferative Diabetic Retinopathy and Vitreous Hemorrhage	HCC122	Proliferative Diabetic Retinopathy and Vitreous Hemorrhage	HCC122	Proliferative Diabetic Retinopathy and Vitreous Hemorrhage	Eye
		HCC124	Exudative Macular Degeneration	HCC124	Exudative Macular Degeneration	

V12 Model (with 70 HCCs)		V21 Model (with 83 HCCs)		V22 Model (with 79 HCCs)		Category Short Name
HCC	Description	HCC	Description	HCC	Description	
HCC130	Dialysis Status	HCC134	Dialysis Status	HCC134	Dialysis Status	Kidney
HCC131	Renal Failure	HCC135	Acute Renal Failure	HCC135	Acute Renal Failure	
HCC132	Nephritis	HCC136	Chronic Kidney Disease (Stage 5)	HCC136	Chronic Kidney Disease (Stage 5)	
		HCC137	Chronic Kidney Disease, Severe (Stage 4)	HCC137	Chronic Kidney Disease, Severe (Stage 4)	
		HCC138	Chronic Kidney Disease, Moderate (Stage 3)			
		HCC139	Chronic Kidney Disease, Mild or Unspecified (Stages 1-2 or Unspecified)			
		HCC140	Unspecified Renal Failure			
HCC148	Decubitus Ulcer of Skin	HCC157	Pressure Ulcer of Skin with Necrosis Through to Muscle, Tendon, or Bone	HCC157	Pressure Ulcer of Skin with Necrosis Through to Muscle, Tendon, or Bone	Skin
HCC149	Chronic Ulcer of Skin, Except Decubitus	HCC158	Pressure Ulcer of Skin with Full Thickness Skin Loss	HCC158	Pressure Ulcer of Skin with Full Thickness Skin Loss	
HCC150	Extensive Third-Degree Burns	HCC161	Chronic Ulcer of Skin, Except Pressure	HCC161	Chronic Ulcer of Skin, Except Pressure	
		HCC162	Severe Skin Burn or Condition	HCC162	Severe Skin Burn or Condition	
HCC154	Severe Head Injury	HCC166	Severe Head Injury	HCC166	Severe Head Injury	Injury
HCC155	Major Head Injury	HCC167	Major Head Injury	HCC167	Major Head Injury	
HCC157	Vertebral Fractures w/o Spinal Cord Injury	HCC169	Vertebral Fractures without Spinal Cord Injury	HCC169	Vertebral Fractures without Spinal Cord Injury	
HCC158	Hip Fracture/Dislocation	HCC170	Hip Fracture/Dislocation	HCC170	Hip Fracture/Dislocation	
HCC161	Traumatic Amputation	HCC173	Traumatic Amputations and Complications	HCC173	Traumatic Amputations and Complications	
HCC164	Major Complications of Medical Care and Trauma	HCC176	Complications of Specified Implanted Device or Graft	HCC176	Complications of Specified Implanted Device or Graft	Complications
HCC174	Major Organ Transplant Status	HCC186	Major Organ Transplant or Replacement Status	HCC186	Major Organ Transplant or Replacement Status	Transplant

V12 Model (with 70 HCCs)		V21 Model (with 83 HCCs)		V22 Model (with 79 HCCs)		Category Short Name
HCC	Description	HCC	Description	HCC	Description	
HCC176	Artificial Openings for Feeding or Elimination	HCC188	Artificial Openings for Feeding or Elimination	HCC188	Artificial Openings for Feeding or Elimination	Openings
HCC177	Amputation Status, Lower Limb/Amputation Complications	HCC189	Amputation Status, Lower Limb/Amputation Complications	HCC189	Amputation Status, Lower Limb/Amputation Complications	Amputation